NUMERICAL TESTING OF PARAMETERIZATION SCHEMES FOR SOLVING PARAMETER ESTIMATION PROBLEMS

L. Velázquez*, M. Argáez and C. Quintero
The University of Texas at El Paso
El Paso, TX 79968-0514

H. Klie and M. F. Wheeler
The University of Texas at Austin
Austin TX 78712

ABSTRACT

We present the numerical performance of two parameterization schemes, Singular Value Decomposition (SVD) and wavelets, for solving automated parameter estimation problems using the Simultaneous Perturbation Stochastic Approximation (SPSA) algorithm. The two schemes are tested on a suit of two large scale two-phase flow problems that illustrate potential for addressing large-scale EQM inverse problems using high-performance computing (HPC).

1. INTRODUCTION

In this paper we present the numerical performance of three parameterization approaches, SVD, wavelets, and the combination of wavelet-SVD for solving automated parameter estimation problems based on the SPSA described in previous reports of this project ([7, 8]). In brief terms, the parameterization methods are based on the principle of projecting the original parameter space onto a lower-dimensional space. In most cases, these projections are computed in terms of SVD (for non-symmetric and rectangular operators), Krylov subspace methods, fast Fourier and wavelet transforms, to name a few alternatives. We conducted the numerical experiments comparing SPSA by using the aforementioned parameterization schemes. It will be shown that the SVD using 50% of the singular values and wavelet level 3 performed extremely well on two test cases of 128x128 gridblocks: channelized (structured) and random (non-structured) permeability fields. We now demonstrate its capabilities for performing parameter estimation using a HPC platform.

2 PROBLEM FORMULATION

A general parameter estimation problem can be written as a nonlinear least squares problem

$$f(x) = \left(G(x) - d\right)^T C_d^{-1} \left(G(x) - d\right). \qquad (1)$$

The first term measures the mismatch between the simulated $G(x)$ and the observed data d, and the observation covariance matrix C_d represents the errors in the data. In our particular case, we assume the matrix to be the identity matrix, and that the model parameter x is the permeability, which is dependent on space. The vector of measurements d may be obtained at different locations and time intervals. In our particular case, we assume that those measurements are pressures observations for the stationary case. We use SPSA, see (Argaez et al., 2007; Klie et al., 2006; Quintero, 2007) to optimize Problem (1).

3 PARAMETERIZATION METHODS

In order to reduce the number of parameters being optimized, we project the original parameter space onto a lower-dimensional space using a parameterization scheme. We consider two schemes: SVD and wavelets. For completeness, we provide a brief description of the parameterization methods being used with SPSA.

3.1 SVD

In this method, we assume that the vector of model parameters x is given by a 2-D permeability field $K \in \Re^{nh \times nv}$. Moreover, we define $x = \log(K)$ to avoid high-local variations of permeability. By applying the SVD approach, x can be decomposed as follow:

$$x = U \Sigma V^t = \sum_{i=1}^{r} \sigma_i u_i v_i^t = \sum_{i=1}^{r} \sigma_i x_i,$$

with $U \in \Re^{nh \times nh}$, the columns of which are composed by the horizontal covariance matrix xx^T, $V \in \Re^{nv \times nv}$, the columns of which are composed by the vertical covariance matrix $x^T x$, and $\Sigma \in \Re^{nh \times nv}$, is a rectangular

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **DEC 2008**	2. REPORT TYPE **N/A**	3. DATES COVERED **-**

4. TITLE AND SUBTITLE **Numerical Testing Of Parameterization Schemes For Solving Parameter Estimation Problems**	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **The University of Texas at El Paso El Paso, TX 79968-0514**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR’S ACRONYM(S)
	11. SPONSOR/MONITOR’S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES
See also ADM002187. Proceedings of the Army Science Conference (26th) Held in Orlando, Florida on 1-4 December 2008, The original document contains color images.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**UU**	**8**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

matrix with diagonal entries $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$, where $r = \min(nh, nv)$ and σ_i is the ith singular value of x.

Let x_i be the ith eigen-image of x that forms an orthogonal basis. Clearly, the contribution of the ith eigen-image to the construction of x depends on the magnitude of σ_i. The error that results from the partial reconstruction can be shown to be equal to the summation of the discarded singular values, thus, providing an indicator of the number of eigen-images required to reconstruct the original field within a prescribed threshold error. This approach has been successfully applied in several permeability estimation problems (Argaez et al., 2007; Banchs et al., 2006; Klie et al., 2006) .

The SVD parameterization boils down to finding the singular values that control the relevant scales of each eigen-image into the estimation. In this work we analyze this strategy by using a 50% of the total number of singular values.

3.2 Wavelet

Wavelet transforms have been used in many subsurface applications showing important potentials for multiscale parameter estimation (Argaez et al., 2007; Rodriguez et al., 2004, 2006). The basic idea is to separate the parameter space representation into distinct frequency packets that are localized in the space or time domain. The parameter space can be conveniently separated in different scales at a low computational cost since wavelets have a compact support (eg, see Liu, 1993; Sahni, and Horne,2006).

Mathematically, the wavelet transform is a method that projects a function *f(x)* onto various vector spaces that represent different scales. Suppose that this function represents a continuous distribution of a particular hydrological property (e.g., permeability or porosity). The measurement of these properties can be realized as

$$F(x) = \int_{-\infty}^{\infty} \Theta(t) f(t-x)\, dt, \tag{3}$$

where $\Theta(x)$, is an averaging kernel such that

$$\int_{-\infty}^{\infty} \Theta(x) dx = 1. \tag{4}$$

This kernel function is a kind of moving average function that is zero outside the region of interest. Thus, different local characteristics of *f(x)* can be defined by choosing a suitable kernel function. Moreover, any scaling of *f(x)* can be obtained by scaling the kernel function. In this sense the kernel function $\Theta(x)$ acts as a low-pass filter or smoothing function. In a few words, the projection of *f(x)* occurs by successive approximation of *f(x)* through the function $\Theta(x)$.

Additionally, details loss between two successive scales can be captured using the so-called wavelet function $\Psi(x)$. Computing inner products *f*(x), $\Theta(x)$ and $\Psi(x)$ and proceeding in a recursive fashion, *f(x)* can then be represented at any desired scale. In this report, we work with resolution scale of wavelet level 3.

3.2 SPSA

SPSA is based on a highly efficient gradient approximation that uses only two function measurements regardless of the dimension of the gradient vector, and it is able to find a good approximation to the solution using few function values (Spall, 1992, 1998). Its disadvantage is that once we have a good approximation, it may not satisfy some conditions and constraints associated with the problem. In particular, SPSA has attracted considerable attention for challenging optimization problems where it is impossible to directly obtain the gradient of the objective function (Rodriguez, 2004, 2007). We use SPSA to find a solution of (1) in combination with the parameterization scheme.

4 Large Scale Two-Phase Flow Problems

We created two test problems of permeability fields as training images for the three parameterization schemes: channelized and random contrast. Our goal is to estimate a permeability field $x \epsilon R^{128x128}$ that computes a pressure data *G(x)* using a Matlab simulator that approximates a given pressure data *d.* The permeability measures the rock's ability to transmit a single fluid at certain conditions. We redefine x=log(x) to avoid high-local variations of permeability. Figures 1 and 2 show the true and priori log permeability and pressure fields of a channelized and random contrast test problems.

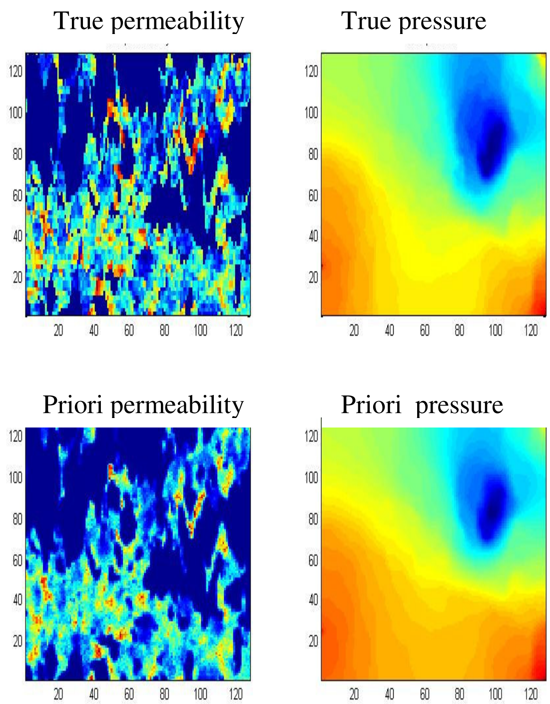

Figure 1. The true (top) and priori (bottom) log permeability and pressure fields of the channelized problem.

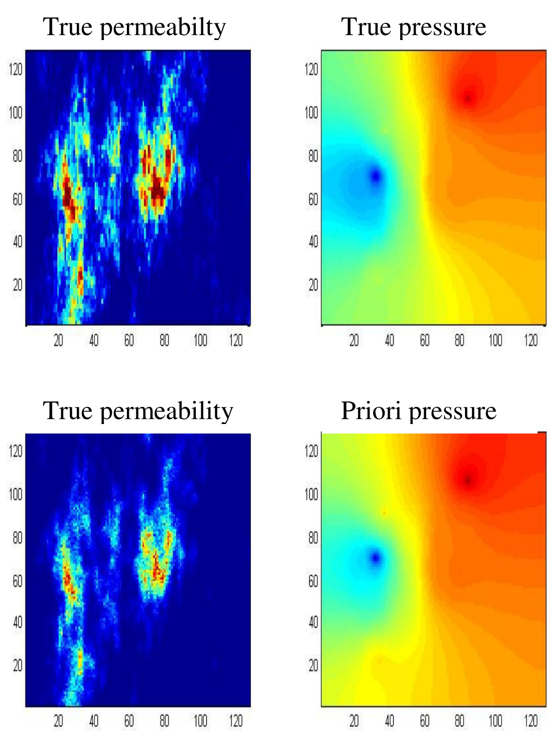

Figure 2. The true (top) and priori (bottom) log permeability and pressure fields of the random problem.

4.1 Sequential Case

We ran SPSA Matlab version 13 in a Dell Laptop, and tabulate the best numerical results obtained for each method: SPSA (non-parameterization), SPSA with parameterization SVD with 50% of singular values, and parameterized SPSA with wavelet. We use 5 initial random points and allow a maximum of 500 and 1000 iterations for SPSA to find a solution for the channelized and random problems, respectively.

In Tables 1-2 we compare the numerical results obtained for the two test problems when using the three parameterization schemes with its corresponding variants in conjunction with the SPSA method.

For each table, the first column states the method being used with SPSA: no parameterization, SVD parameterization with 50% of singular values being used), Wavelet parameterization level 3, and Wavelet-SVD (wavelet with 3 level combined with SVD using a 50% of the singular values). We should remark that the wavelet level represent the resolutions of 16x16 (i.e. 64 parameters to be optimized) for the given 128x128 data. The next three columns tabulates the parameter space, the number of SPSA iterations and function evaluations. The fifth column indicates the relative error R_P between the predicted and observed data and it is given by:

$$R_P=\|F_P - F_O\|/\|F_O\|,$$

where F_P and F_O denote the number of function evaluations performed by the parameterized SPSA and non-parameterized SPSA, respectively. This is one of the metrics proposed in this paper. The last column indicates the ratio of the total number of function evaluations between the parameterized and non-parameterized SPSA versions (F_P/F_O).

We run the original SPSA until either $R_P \leq 0.05$ or the maximum number of SPSA iterations have been achieved (recall that there are 3 function evaluations in each SPSA iteration), then this value of R_P becomes the target or stopping criteria for the parameterization methods. We expect the parameterized version will take less iterations due to the number of reduced parameters being involved for solving the problems.

Table 1. Numerical results generated for the channelized problem.

Method	n	Iter	#F	R_P	F_P/F_O
SPSA	16384	550	1650	0.0510	1.0000
SVD (50%)	64	41	123	0.0501	0.0745
Wavelet 3	256	132	396	0.0503	0.2400
Wavelet – SVD (50%)	8	312	104	0.0552	0.1891

Table 2. Numerical results generated for the random problem.

Method	n	Iter	#F	R_P	F_P/F_O
SPSA	16384	1000	3000	0.1045	1.0000
SVD (50%)	64	58	174	0.1042	0.0580
Wavelet 3	256	87	261	0.1024	0.0870
Wavelet –SVD (50%)	8	216	72	0.1037	0.0720

Table 1 shows that the best strategy for the channelized problem is the parameterized method SVD with 50% of the singular values (64 parameters optimized) that gives less number of iterations, and lowest value R_P, and the second option is Wavelet. In the case of the wavelet-SVD parameterization schemes did not converge after 100 iterations ($R_p \geq 0.05$). In the case of the random problem, the option wavelet gave the lowest R_P value but in terms of function evaluations is better SVD with 50% with a competitive R_P value.

Figures 3-4 show the history of the function values at each iteration obtained for each of the test problems using the original SPSA and the parameterized versions using SVD, wavelet, and wavelet-SVD options. We observe that for both problems the non-parameterized SPSA takes considerable more iterations than the parameterized SPSA methods.

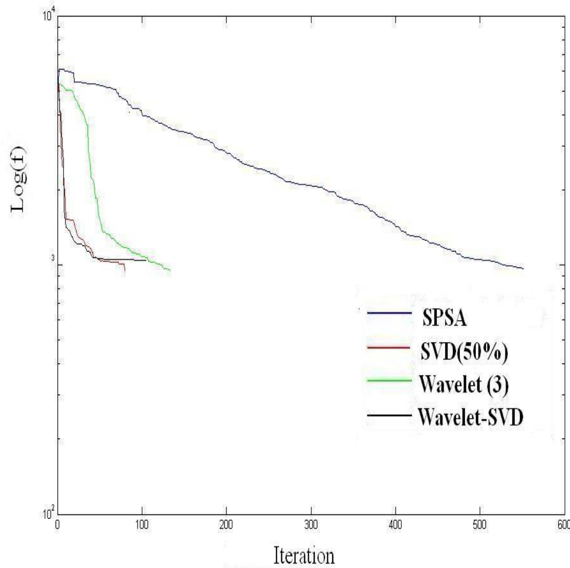

Figure 3. Channelized problem.

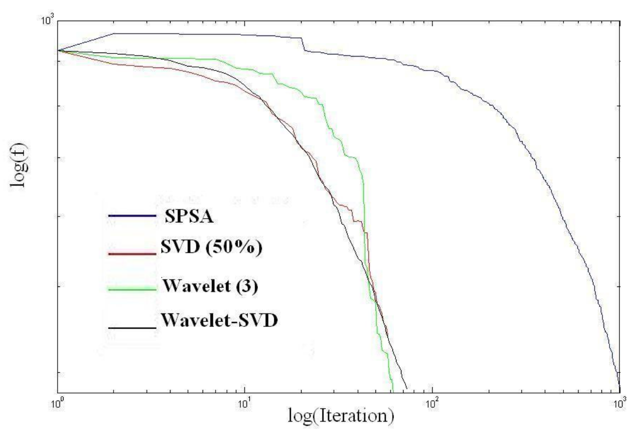

Figure 4. Random problem.

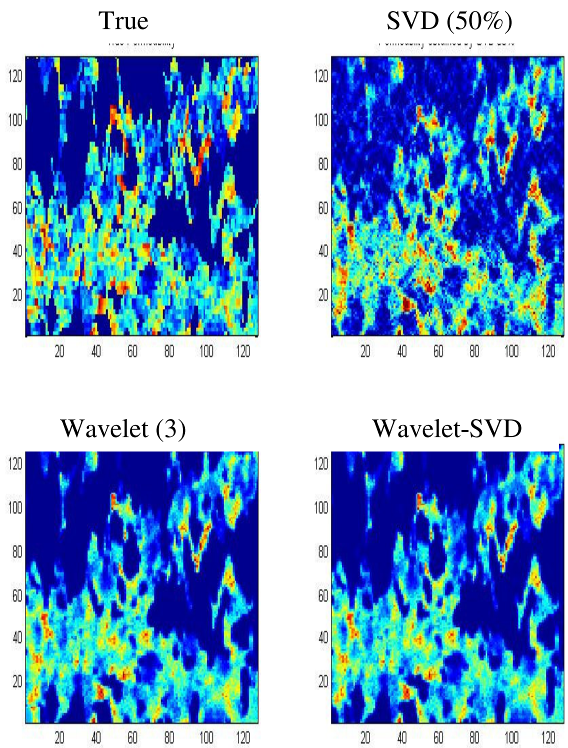

Figure 5. Numerical results of the best permeability fields obtained for the channelized problem.

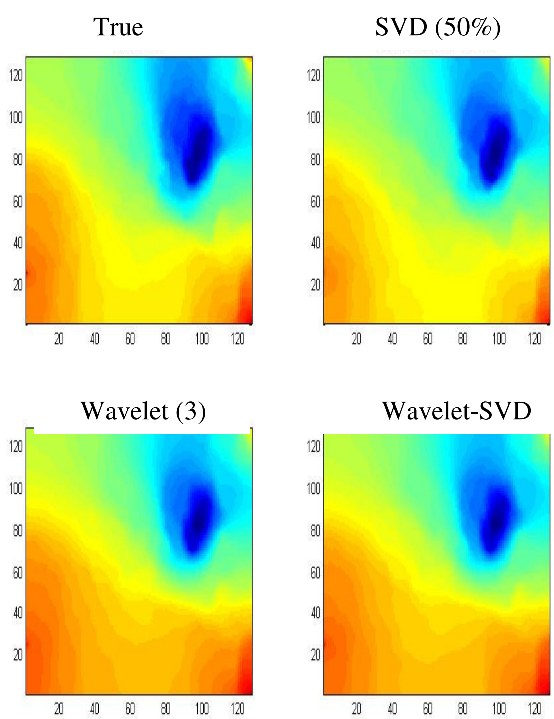

Figure 6. Numerical Results of the best pressure fields obtained for the channelized problem.

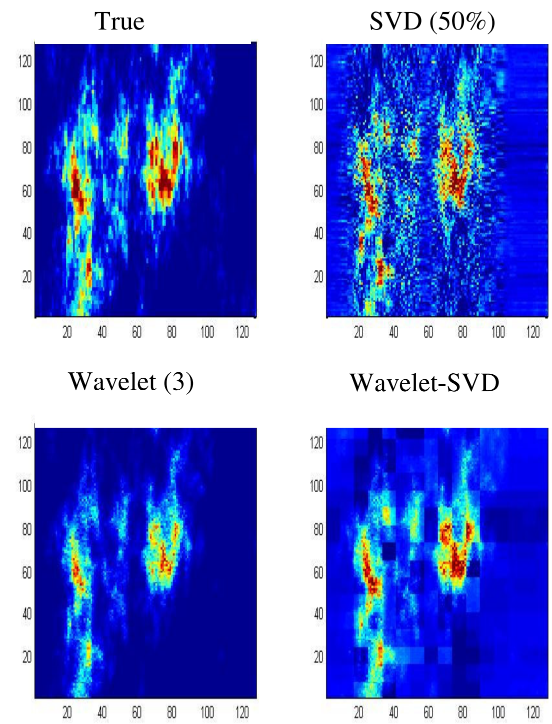

Figure 7. Numerical results of the best permeability fields obtained for the random problem.

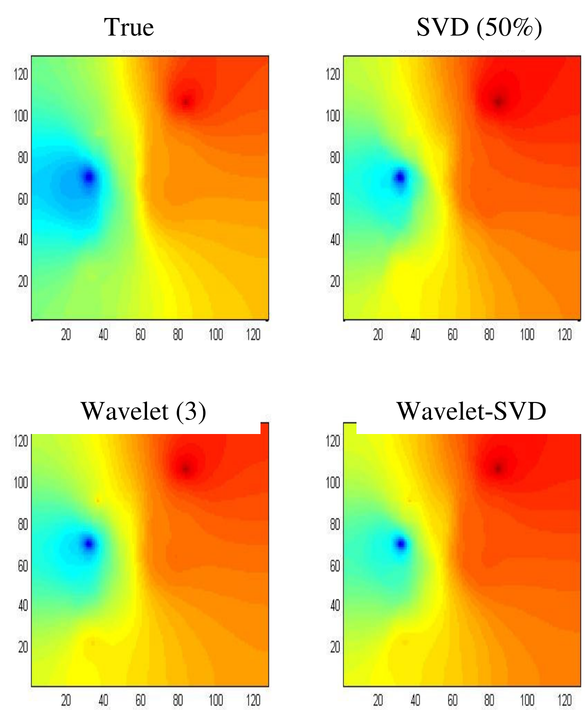

Figure 8. Numerical Results of the best pressure fields obtained for random problem

The results obtained for the two test problems were promising, therefore a numerical experimentation using HPC was considered in this final report.

5. HPC Numerical Experimentation

We run the each of the parameterization schemes with SPSA in conjunction with the Matlab simulator framework IPARS (Integrated Parallel Accurate Reservoir Simulator) (Rodriguez et al., 2004, 2007) on a UTEP machine Virgo. The problem consists of finding a permeability field that involves 16,394 parameters. The field is parameterized by SVD (Rodriguez et al., 2007) reducing the original parameter space to only 64. In the case of using wavelet level 3, it reduces to only 256 parameters.

The following figure illustrates the optimization scheme used to solve the two large scale problems.

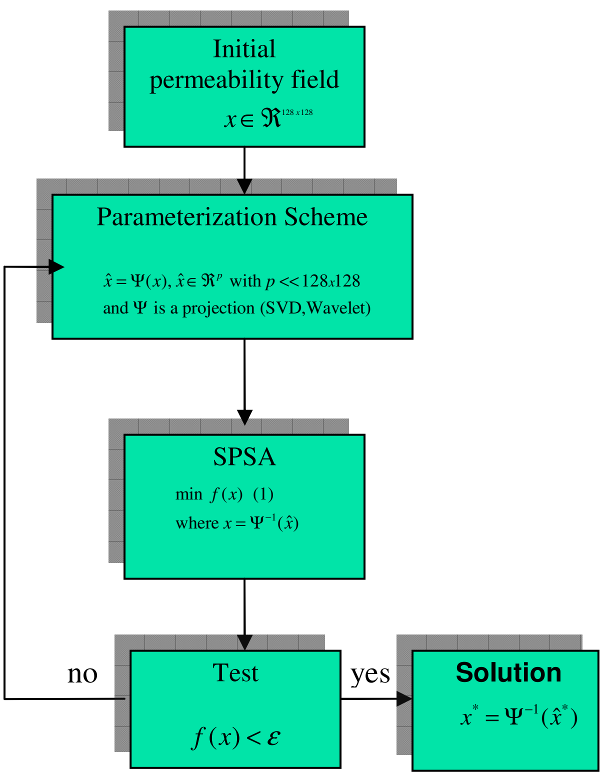

Fig. 9. Flowchart of the optimization procedure.

The UTEP's VIRGO machine is a Beowulf cluster that employs a front end and 20 compute nodes, each with 2 Intel Xeon 3.06 GHz processors and 4 GB main memory, 487 GB user disk space and a 139 peak Gflops rating. This machine has the Matlab and Wavelet toolbox installed on each of the 20 processors. We only had access to up 16 processors, therefore the experimentation was done using 1, 2, 4, 8, and 16 processors. We did not use a machine with 256-512 processors to conduct our experimentation because the Matlab Wavelet toolbox was not available at the machine. Such requirement was needed to be bought for each of the processors, and the funds were not available.

For each problem, we run the following methods: non-parameterization, SVD with 50% of singular values and Wavelet level 3. SPSA stopped if the maximum number of function evaluation of 500 was reached, i.e. only up to 166 iterations were allowed. We run a multistart SPSA on 1, 2, 4, 8 and 16 processors. A multistart SPSA consist in starting SPSA 16 times, and dividing the work in the number of processor available. We use the same initial point for each parameterization scheme, but SPSA uses a different random seed each time, so the final solution is different.

In Tables 3-8 we tabulate the numerical results obtained for the two test problems when using non-parameterization and the two parameterization schemes with the SPSA method. We should remark that the SVD parameterization scheme was run using 50% of the singular values, and the wavelet level 3 represents the 16x16 resolution for the given 128x128 data.

For each table, the first column represents the number of processor P being used with SPSA to solve the problem, and the second column states the mean time T_P taken to solve the problem. The third column indicates the mean relative error R_P. The fourth column represents the speed up S_P, and we rely on at least 16 processors to achieve a sustained

$$S_P = \frac{T_1(F,1)}{T_P(P*F,P)} \geq 0.5,$$

where F is the number of function evaluations for a single processor and $T_P(W,P)$ is the total wall clock time used for solving the optimization problem with W function evaluations on P processors. The last column indicates the efficiency defined by $E_P=S_P/P$.

Table 3. Non-parameterization results for the channelized problem.

P	T_P	R_P	S_P	E_P
1	4165.100	0.58352	1.000	1.000
2	2097.800	0.56944	1.985	0.993
4	1048.400	0.56128	3.973	0.993
8	530.100	0.58907	7.857	0.982
16	276.200	0.5955	15.080	0.943

Table 4. SVD parameterization results for the channelized problem.

P	T_P	R_P	S_P	E_P
1	4072.300	0.1824	1.000	1.000
2	3714.400	0.1868	1.096	0.548
4	1025.800	0.1856	3.970	0.992
8	516.900	0.1847	7.878	0.985
16	260.300	0.1841	15.644	0.978

Table 5. Wavelet parameterization results for the channelized problem.

P	T_P	R_P	S_P	E_P
1	4548.400	0.27249	1.000	1.000
2	2276.700	0.28680	1.998	0.999
4	1136.600	0.28608	4.002	1.000
8	577.100	0.28116	7.882	0.985
16	290.800	0.26419	15.639	0.977

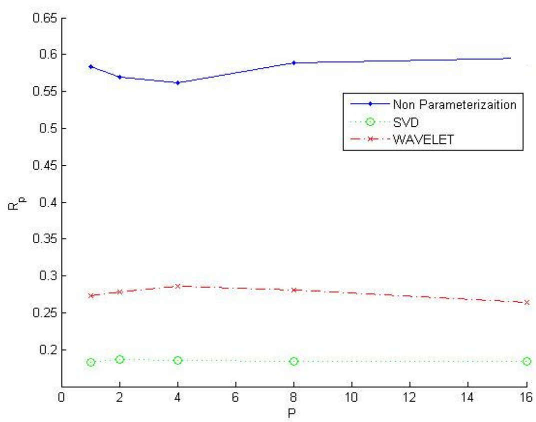

Fig. 10. Numerical results of the mean error R_P using different processors P obtained for the channelized problem. Parameterizations

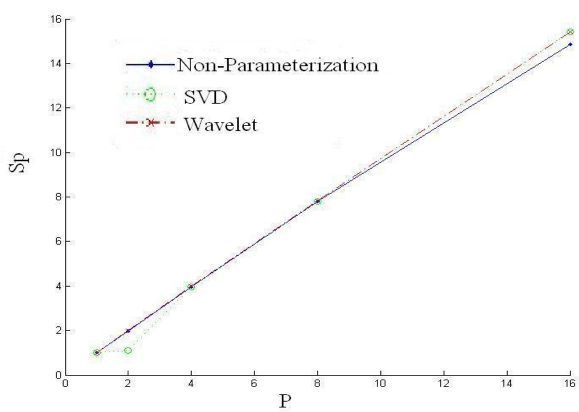

Figure 11. Numerical results of the speed up S_P using different processors P obtained for the channelized problem via non-parameterization, SVD and Wavelet parameterizations.

Table 6. Non-parameterization results for the random problem.

P	T_P	R_P	S_P	E_P
1	8522.300	0.815973	1.000	1.000
2	4299.200	0.775318	1.982	0.991
4	2244.800	0.787177	3.796	0.949
8	1108.100	0.801091	7.691	0.961
16	618.100	0.802728	13.788	0.862

Table 7. SVD parameterization results for the random problem.

P	T_P	R_P	S_P	E_P
1	8191.700	0.511193	1.000	1.000
2	4108.300	0.521353	1.994	0.997
4	2110.800	0.524534	3.881	0.970
8	1108.600	0.538929	7.390	0.924
16	576.500	0.524443	14.210	0.888

Table 8. Wavelet parameterization results for the random problem.

P	T_P	R_P	S_P	E_P
1	9014.700	0.500469	1.000	1.000
2	4428.500	0.491750	2.036	1.018
4	2280.800	0.527515	3.952	0.988
8	1141.400	0.556861	7.898	0.987
16	656.600	0.584490	13.729	0.858

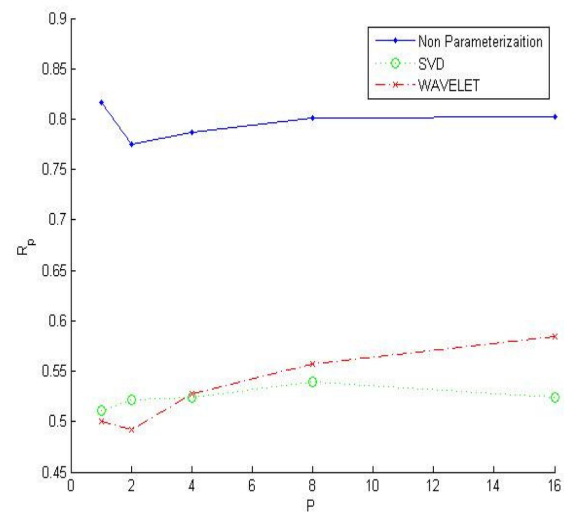

Fig. 12. Numerical results of error R_P using different processors P obtained for the random problem via non-parameterization, SVD and Wavelet parameterizations.

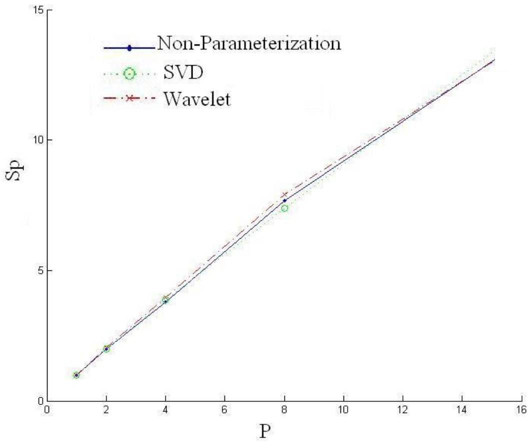

Fig. 13. Numerical results of speed up S_P using different processors P obtained for the random problem via non-parameterization, SVD and Wavelet parameterizations.

6. CONCLUSIONS

We have tested three parameterization methods to mitigate the curse of dimensionality that typically arises in many Department of Defense (DoD) parameter estimation scenarios. We have made a numerical assessment of SVD with 50% of singular values, wavelet level 3 and hybrid wavelet-SVD parameterization schemes. The main challenge of the parameterization methods is to be able to capture all possible features from the parameter space into a lower dimensional space representation.

Our numerical results show that the proposed methods converge in significantly reduced number off iterations for a channelized and a random permeability field. In view of this, it is highly suggested to the DoD users to incorporate some of these methods in their large-scale parameter estimation work in order to increase efficiency and accuracy of subsurface property estimations that are subject to the curse of dimensionality.

ACKNOWLEDGEMENTS

The authors thank DoD for the support given with the grant PET Project EQM-KY7-003. The UT-El Paso team also acknowledges the support of the NSF Cyber-Share HRD-0734825 and the Department of the Army ARL Grant No. W911NF-07-02-0027.

REFERENCES

Argaez, M., Quintero, C., Klie, H., M.F. Wheeler, and Velazquez, L., 2007: Hybrid Optimization Approach for Solving Parameter Estimation Problems, Proc.,*6th International Computational and Industrial Applied Mathematics*, Zurich, SW.

Banchs,R., Klie, H., Rodriguez, A., and Wheeler, M.F.. 2006: A Neural Stochastic Optimization Framework for Oil Parameter Estimation, *Intelligent Data Engineering and Automated Learning (IDEAL)*, Lecture Notes in Computer Science, 147-154.

Klie, H., Rodríguez, A., Banchs, R., and Wheeler, M.F., 2006: Assessing the Value of Sensor Information in 4-D Seismic History Matching, Proc., *76th SEG International Exposition & Annual Meeting*, New Orleans, LA.

Liu, J., 1993: *A multiresolution method for distributed parameter-estimation*, Siam Journal on Scientific Computing **14**, no. 2, 389-405.

Quintero, C., 2007: A Hybrid Optimization Approach for Solving Parameter Estimation Problems, The University of Texas at El Paso, 1-56.

Rodriguez, A., Klie, H., Thomas, S. G. and Wheeler, M. F., 2004: A Multiscale and Metamodel Simulation-Based Method for History Matching, Proc.,*10th European Conference on the Mathematics of Oil Recovery*, Euroepan Association of Geoscience and Engineers, Amsterdam, ND.

Rodriguez, A., Klie, H., Wheeler, M.F. and Banchs, R.,2007: *Assessing multiple resolution scales in history matching with stochastic neural metamodels*, SPE Reservoir Simulation Symposium.

Sahni, I. and Horne, R. N., 2006: *Stochastic history matching and data integration for complex reservoirs using a wavelet-based algorithm,* Proc., SPE Reservoir Simulation Symposium, Houston, TX, SPE #105824.

Spall, J.C, 1992: Multivariate Stochastic Approximation Using a Simultaneous Perturbation Gradient Approximation, *IEEE Transactions on Automatic Control* **37**, 332-341.

Spall, J.C., 1998: Implementation of the Simultaneous Algorithm for Stochastic Optimization, *IEEE Transactions on Aerospace and Electronic Systems* **34**, 817-823.